

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : <b>C12N 15/70, 15/62, 9/24, 1/21, A61K 38/47, C12Q 1/34 // (C12N 1/21, C12R 1:19)</b>		A1	(11) International Publication Number: <b>WO 97/02351</b> (43) International Publication Date: <b>23 January 1997 (23.01.97)</b>
(21) International Application Number: <b>PCT/GB96/01577</b> (22) International Filing Date: <b>1 July 1996 (01.07.96)</b>  (30) Priority Data: <b>9513683.4 5 July 1995 (05.07.95) GB</b>		(81) Designated States: AL, AU, BB, BG, BR, CA, CN, CZ, EE, GE, HU, IL, IS, JP, KP, KR, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  Published <i>With international search report.</i>	
(71) Applicants ( <i>for all designated States except US</i> ): CIBA-GEIGY AG [CH/CH]; Klybeckstrasse 141, CH-4002 Basle (CH). CAMBRIDGE UNIVERSITY TECHNICAL SERVICES LIMITED [GB/GB]; The Old Schools, Trinity Lane, Cambridge CB2 1TS (GB).  (72) Inventors; and (75) Inventors/Applicants ( <i>for US only</i> ): TAYLOR, Peter, William [GB/GB]; Merringdean Oak, Merringdean Road, Billingshurst, West Sussex RH14 9HF (GB). LUZIO, John, Paul [GB/GB]; Pippin Meadow, Lowfields, Little Eversden, Cambridge CB3 7HJ (GB). BRYANT, Jonathan, Mark [GB/GB]; 23 Monmouth Street, Topsham, Exeter EX3 0AJ (GB).  (74) Agent: SHARMAN, Thomas; Ciba-Geigy PLC, Patent Dept., Hulley Road, Macclesfield SK10 2NX (GB).			
(54) Title: RECOMBINANT PROTEIN HAVING BACTERIOPHAGE ENDOSIALIDASE ENZYMATIC ACTIVITY			
(57) Abstract <p>A recombinant protein having bacteriophage endosialidase enzymatic activity obtainable by expression from a recombinant vector comprising a DNA sequence encoding a bacteriophage endosialidase linked to a DNA sequence of an expression vector which expresses a polypeptide which adds to the N-terminus of the endosialidase, or an analogue of said protein, which is a mutant, functional fragment or derivative of said protein having endosialidase enzymatic activity.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Iceland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EZ	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

RECOMBINANT PROTEIN HAVING BACTERIOPHAGE ENDOSIALIDASE ENZYMATIC ACTIVITY

This invention relates to a recombinant protein having bacteriophage endosialidase activity, to a process for the production thereof and to recombinant expression systems for use in the production thereof.

Bacteriophage E is a member of the PK1A-PK1E family of phages; these phages were isolated originally from European sewage to aid in the clinical identification of Escherichia coli K1 infections, which can result in high mortality rates in cases of neonatal meningitis. Bacteriophage E endosialidase (K1E endosialidase) is thought to be the enzyme responsible for initial binding to host bacteria by specifically recognising and hydrolysing the  $\alpha$ -2,8-linked poly-N-acetylneurameric acid (polysialic acid/PSA) carbohydrate polymers of the K1 glycocalyx.  $\alpha$ -2,8-linked PSA is also expressed on the cell surface of several other pathogenic bacteria, and various tumour cells and cell lines. It has been proposed in U.S. Patent No. 4 695 541 that K1E endosialidase could be used in the diagnosis and therapy of K1 meningitis, septicaemia or bacteraemia due to the enzyme's high specificity for hydrolysing  $\alpha$ -2,8-sialosyl linkages. PSA has been suggested as an oncodevelopmental marker in human tumours of the kidney and neuroendocrine tissues and also may contribute to the invasive and metastatic potential of some tumours.

In J. Bacteriol., 1993, 175, 4354 - 4363, there are described attempts to obtain enzymatically active protein by expression from a DNA construct derived from the related K1F phage; these attempts were unsuccessful.

It has now been found that protein having bacteriophage endosialidase enzymatic activity, i.e. a protein which specifically binds to or cleaves  $\alpha$ -2,8-polysialic acid, can be obtained by expression from a DNA construct which is derivable from the K1E endosialidase gene and is cloned into an expression vector which expresses a polypeptide which adds to the N-terminus of the endosialidase sequence.

- 2 -

Accordingly, the present invention provides, in one aspect, a recombinant protein having bacteriophage endosialidase enzymatic activity obtainable by expression from a recombinant vector comprising a DNA sequence encoding a bacteriophage endosialidase linked to a DNA sequence of an expression vector which expresses a polypeptide which adds to the N- terminus of the endosialidase, or an analogue of said protein which is a mutant, functional fragment or derivative of said protein having endosialidase enzymatic activity.

The mutant may be, for example, a protein having an amino acid substituted or deleted at one or more positions. The functional fragment may be C- or N- terminal shortened fragment or a fragment from within the polypeptide chain which has endosialidase enzymatic activity. The derivative may be, for example, a pharmaceutically acceptable salt with an acid such as hydrochloric acid, sulphuric acid, phosphoric acid, pyrophosphoric acid, benzenesulphonic acid, p-toluenesulphonic acid, methanesulphonic acid, lactic acid, palmic acid, tartaric acid, ascorbic acid, or citric acid; with a base, usually a nitrogen containing base such as sodium, potassium, magnesium or ammonium nitrogen-containing base; or an internal salt.

In another aspect, the present invention provides a recombinant vector comprising a DNA sequence encoding a bacteriophage endosialidase linked to a DNA sequence of an expression vector which expresses a polypeptide which adds to the N- terminus of the endosialidase, said recombinant vector being capable of directing expression of said protein in a compatible host cell.

In a further aspect the present invention provides a process for the production of a protein having bacteriophage E endosialidase enzymatic activity which comprises culturing a host cell transformed with a recombinant vector as hereinbefore defined under conditions allowing expression of said protein and isolating the protein thereby produced. In a yet further aspect, the present invention provides a host cell transformed with a recombinant vector as hereinbefore defined.

Preferred protein according to the invention is a protein obtainable by expression from a recombinant vector as hereinbefore defined in which the DNA sequence encoding the endosialidase is derived from a DNA construct encoding amino acid residues encoded by nucleotides I72 to I744 of the bacteriophage E endosialidase gene, i.e. nucleotides I72 to I744 of SEQ ID No. I as hereinafter defined, or a mutant, functional fragment or derivative of said protein which has endosialidase enzymatic activity. An especially preferred protein according to the invention is a protein obtainable by expression from a recombinant vector as hereinbefore defined in which the DNA sequence encoding the endosialidase is derived from a DNA construct encoding amino acid residues encoded by nucleotides I to 2436 of the bacteriophage E endosialidase gene, i.e. nucleotides I to 2436 of SEQ. ID NO. I as hereinafter defined, or a mutant, functional fragment or derivative of said protein having endosialidase enzymatic activity.

The protein of the invention is generally expressed in the form of a fusion protein comprising the endosialidase linked, directly or through a spacer, to a polypeptide derived from the expression vector, i.e. the vector used for expression of the protein in a suitable host cell, a preferred such polypeptide being glutathione S-transferase. Where it is desired that the polypeptide components of the fusion protein should be separable, if the fusion protein does not naturally contain a region which can be specifically cleaved chemically or enzymatically, such a region can be inserted using conventional procedures. Examples of selective cleaving reagents or cleaving enzymes for fusion proteins are V8 protease, trypsin, thrombin, factor X, CNBr, peptidase *yscα* and *yscF*.

In a particularly preferred embodiment of the invention, the protein of the invention is in the form of a fusion protein comprising bacteriophage E endosialidase linked to glutathione S-transferase, the fusion protein preferably having a molecular weight of about 100 kDa.

A DNA construct, i.e. recombinant DNA molecule, suitable for the expression of a protein according to the invention may be an isolated DNA fragment encoding a bacteriophage endosialidase, for example consisting only of the coding region or prolonged by homologous or heterologous DNA sequences. The construct may be a DNA fragment encoding the endosialidase cloned into a suitable cloning vector, preferably a bacterial

vector such as pBR317, pBR322, pUC18, pSF2124 or, especially, Bluescript SK<sup>+</sup>. Where such a clone lacks convenient restriction sites with which to isolate solely the endosialidase open reading frame, it may be amplified by a polymerase chain reaction (PCR) using primers incorporating the restriction sites required.

The DNA fragment encoding the bacteriophage endosialidase may be obtained from genomic bacteriophage E DNA or a synthetic DNA that is substantially homologous thereto, i.e. is 80-100% homologous thereto. Bacteriophage E can be purified and total genomic DNA can be extracted using conventional procedures. The extracted DNA can then be digested with an appropriate restriction enzyme such as Bgl II, Eco RI, Hinc II, Hind III, Bam HI or Pst I. The digestion product can be subjected to preparative electrophoresis with low-melting point agarose gel to enrich DNA fractions of a certain length in order to enrich DNA fragments encoding the protein of the invention.

When a nucleotide sequence encoding the bacteriophage endosialidase, or an amino acid sequence thereof, is known, DNA encoding the endosialidase can also be prepared by methods leading directly to the desired DNA such as conventional PCR procedures or in vitro chemical synthesis.

For expression of a protein of the invention, the DNA construct is cloned into an expression vector which expresses polypeptide which adds to the N- terminus of the endosialidase to give a recombinant vector according to the invention. The expression vector is, of course, chosen according to the nature of the host cell chosen for expression of the protein. Suitable such expression vectors are available commercially. Expression is preferably carried out in a prokaryotic host, more preferably a microbial host, especially *E. coli*, when a suitable expression vector is a prokaryotic expression vector such as a phage  $\lambda$  or a bacterial plasmid. Examples of particular prokaryotic expression vectors are pGEX vectors, e.g. pGEX-2T (Pharmacia Biotech), which result in the expression of an endosialidase - glutathione S-transferase (GST) fusion protein, pMAL (New England Biolabs) which results in expression of an endosialidase - maltose binding protein fusion protein, the 'pinpoint' system from Promega which biotinylates expressed protein, the 'strep-tag' system from Biometra which places a streptavidin binding peptide on expressed protein, the Ni-NTA

system from Qiagen which adds 6 histidines to expressed protein to bind nickel, and the Xpress system from Invitrogen working on a similar principle to the Ni-NTA system.

Preferred expression vectors are pGEX vectors, which have a tac promoter, an internal lac I<sup>r</sup> gene and a thrombin or factor X<sub>a</sub> protease recognition site, especially pGEX-2T which has the sequence:

Leu Val Pro Arg Gly Ser Pro Gly Ile His Arg Asp  
CTG GTT CCG CGT GGA TCC CCG GGA ATT CAT CGT GAC TGA CTG ACG

Cloning of the DNA construct into the expression vector to give the recombinant vector of the invention may be carried out using conventional restriction and ligation techniques. Thus, where the DNA construct contains Bam HI and EcoRI restriction sites, which may have been incorporated by PCR amplification, the DNA construct and the expression vector may be digested simultaneously with Bam HI and EcoRI and ligation effected using a DNA ligase in accordance with the manufacturer's instructions.

As mentioned hereinbefore, the host cells used for expression of a protein of the invention are preferably prokaryotic, more preferably microbial cells, including cells of bacteria such as *Bacillus subtilis*, *Pseudomonas*, *Streptococcus* or, especially, *E. coli*.

Transformation of the host cells may be carried out using conventional techniques appropriate for those cells. Accordingly, the transformation procedure for *E. coli* cells includes, for example, Ca<sup>2+</sup> pretreatment of the cells so as to allow DNA uptake, and incubation with the recombinant vector. The subsequent selection of the transformed cells can be achieved, for example, by transferring the cells to a selective growth medium which allows separation of the transformed cells from the parent cells, or by restriction analysis of a miniprep DNA sample obtained from the incubated cells.

The transformed host cells may be cultured by methods known in the art in a liquid medium containing an assimilable source of carbon, e.g. a carbohydrate such as glucose or lactose, nitrogen, e.g. an amino acid, peptide, protein or degradation product thereof such as a peptone, ammonium salt or the like, and an inorganic salt, e.g. a sulfate, phosphate and/or

carbonate of sodium, potassium, magnesium or calcium. The medium may also contain, for example, a growth-promoting substance, such as a trace element, for example iron, zinc, manganese and the like.

Culturing may be effected by processes which are known in the art. The culture conditions, such as temperature, pH value of the medium and fermentation time, are chosen so that a maximum expression level of the protein of the invention is obtained. Thus, an *E. coli* strain is preferably cultured under aerobic conditions by submerged culture with shaking or stirring at a temperature of about 20°C to 40°C, preferably at about 37°C, and a pH value of 4 to 8, preferably of about 7, for about 4 to 30 hours, preferably until maximum yields of the protein of the invention are reached.

The expressed protein can be extracted from microbial cells such as *E. coli* cells or a supernatant of a cell culture by conventional methods, e.g. comprising lysis of the cells, chromatography such as ion-exchange, hydrophobic or size-exclusion chromatography, precipitation, e.g. with ammonium sulfate or acid, preparative electrophoresis such as sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS-PAGE) or isoelectric focussing, and the like. When, as in especially preferred embodiments of the invention, the expressed protein is an endosialidase-glutathione S-transferase fusion protein, this may be purified by binding to glutathione beads as described by Smith and Johnson (1988) Gene 67, 31-40. Cleavage of the purified fusion protein can be effected with thrombin, for example following the instructions of Pharmacia Biotech, manufacturers of the pGEX-2T expression vector.

The present invention also provides a pharmaceutical composition comprising as active ingredient a protein of the invention or a pharmaceutically acceptable salt thereof, optionally together with a physiologically acceptable carrier, which may be, for example, an excipient, diluent or other conventional auxilliary in pharmaceutical compositions.

Proteins of the invention may be used in the diagnosis or treatment of medical conditions, especially of the human body, including various diseases, particularly meningitis and cancers characterised by expression of polysialic acid on the surface of the tumour cell, such as Wilm's Kidney Tumour, small cell lung carcinoma, neuroblastoma, medullary thyroid

carcinoma, urinary tract tumour, neuroectodermal tumour, teratoma, rhabdomyosarcoma, pheochromocytoma, Ewing's sarcoma, insulinoma, breast cancer and pituitary tumour. The proteins may be used to inhibit tumour metastasis, for example post-surgical metastasis. The proteins may also be used in the diagnosis or treatment of other conditions caused by *E. coli* K1, such as sepsis and urinary tract infections, or by other bacteria expressing polysialic acid on the cell surface thereof.

Thus the present invention also provides a method of treating a condition caused by a bacterium expressing polysialic acid on a cell surface thereof, cancer characterised by expression of polysialic acid on a tumour cell surface, or tumour metastasis, which comprises administering a protein or analogue of the invention as hereinbefore defined to a warm-blooded mammal in need of such treatment.

A pharmaceutical composition of the invention, particularly for the above indications, may be administered parenterally, for example intravenously, intracutaneously, subcutaneously or intramuscularly. The dosage depends principally on the method of administration and on the purpose of the treatment. Individual doses and the administration regime can best be determined by individual judgement of a particular case of illness. Usually, a therapeutically effective amount of a protein of the invention, when administered by injection, is from about 0.005 to about 0.1 mg/ kg body weight.

In addition to the active ingredient, an injectable pharmaceutical composition of the invention may contain a buffer, for example a phosphate buffer, sodium chloride, mannitol or sorbitol to adjust the isotonicity, and an antibacterially active preservative such as the methyl or ethyl ester of p-hydroxybenzoic acid.

The proteins of the invention, in view of their enzymatic activity, may also be used in the analysis of glycoproteins, for example detection and sequencing of oligosaccharide moieties decorating glycoproteins, since they can selectively remove particular sugar residues from the glycoproteins.

The invention is illustrated by the following Examples, which relate to especially preferred embodiments.

Example 1 Preparation of DNA Construct Containing Bacteriophage E Endosialidase Open Reading Frame

Unless otherwise stated, all procedures used are as described by Sambrook et al, Molecular Cloning: a Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, New York (1989).

Degenerate oligonucleotide probes are designed with reference to *E. coli* codon usage tables (Holm (1986) Nuc. Acids Res. 14, 3075-3087), prepared using an automated Applied Biosystems PCR - MATE model 391 DNA synthesiser and 5' end-labelled with [ $\gamma$ -<sup>32</sup>P]ATP (Amersham International Plc., Amersham, Bucks, U.K.) using T4 polynucleotide kinase. The radiolabelled oligonucleotide probes are hybridised to restriction enzyme digests of bacteriophage E DNA, electrophoresed in agarose gels and transferred to Hybond-N nylon membrane (Amersham International Plc.). Bacteriophage E DNA fragments reacting with the probes are identified by autoradiography, purified from NA grade agarose gels (Pharmacia Biosystems Ltd, Milton Keynes, Bucks, U.K.) and ligated into Bluescript SK+ (Stratagene Inc., La Jolla, CA, USA) using T4 DNA ligase (NEB Inc.). Transformations of *E. coli* Epicurian SURE cells (Stratagene Inc.) with Bluescript SK+ are conducted according to an electroporation method (Dower et al (1988) Nucleic Acids Res. 16, 6127 - 6145) using a Bio-Rad Gene Pulser and Pulse Controller, or alternatively high efficiency *E. coli* JM109 competent cells (Promega Inc. Madison, WI, USA) are transformed by heat shock at 42°C for 60 sec. Clones transformed with recombinant plasmid are identified by growing on 2TY /ampicillin agar plates and using a mixture of 50 mg/ml 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-Gal) and 0.1M isopropyl β-D- thiogalactopyranoside (IPTG) to allow blue-white colour selection of colonies. Double stranded DNA sequencing is conducted using the Sequenase Version 2.0 sequencing kit from United States Biochemical Corporation, Cleveland, OH, USA and a model SA sequencing apparatus from BRL Life Technologies Inc., Gaithersburg, MD, U.S.A. Sequencing is facilitated by the technique of nested deletions or by using synthetic oligonucleotide primers prepared by British Biotechnology Products Ltd, Abingdon, Oxon, U.K. or as above.

A degenerate oligonucleotide probe, Probe 1 [5'-TAC(T)CAC(T)CAGGGT(G)GAC (T)GTG(T)GCG(C)CC-3'], is derived from the cyanogen bromide fragment of KIE

endosialidase with the longest unambiguous amino acid sequence, and is the least degenerate of five probes designed using the partial amino acid sequences obtained from the cyanogen bromide fragments. A 1.9 kb BgIII restriction digest fragment of genomic bacteriophage E DNA is identified as potentially encoding endosialidase sequence by Southern blot analysis using  $^{32}$ P-radiolabelled probe 1. BgIII and BamHI restriction endonucleases generate cohesive protruding ends with the same sequence and this enables the ligation of the 1.9kb BgIII fragment into the BamHI site of Bluescript SK+ cloning vector (Promega Inc). Plasmid miniprep DNA from a clone transformed with the resultant recombinant vector (Clone 1) yields DNA sequence which encodes a deduced protein sequence containing a stretch of sequence identical to that of the CNBr fragment used to design Probe 1.

Probe 2 [5'-GATCTTGGTCTAATCCCT-3'], a non-degenerate oligonucleotide 18-mer, is synthesised using the sequence at the 5' end of Clone 1. This probe identifies one of two SmaI digest fragments of genomic bacteriophage E DNA which runs as a singlet equivalent to about 3.3kb. It is verified that this fragment codes for DNA sequence upstream of the 5' end of Clone 1 by digesting the Clone 1 insert DNA with SmaI. The result of this digest shows there are at least 3 SmaI sites in the Clone 1 insert DNA, the largest fragment being 1.1kb. Since restriction analysis of bacteriophage E DNA shows that there are only two BgIII sites in the whole genome, the gel purified SmaI fragments are digested with BgIII and the fragment containing the probe 2 recognition sequence and the BgIII site yields two fragments of 2.1kb and 1.1kb. The 2.1kb SmaI x BgIII digest fragment is cloned into Bluescript SK+ by ligation of the BgIII end to a BamHI end, followed by end-filling using the Klenow fragment of T4 DNA polymerase and ligating the resultant blunt ends together to circularise the plasmid. The resultant clone (Clone 2) is found to contain an open reading frame encoding the N-terminus of KIE endosialidase by comparison with the N-terminal amino acid sequence of the ~ 76 kDa enzyme subunit. Overlapping sequence is obtained for clones 1 and 2 in both 5' and 3' directions, and the positions of open reading frames are determined by codon preference and positional base preferences analysis (Staden et al, (1982) Nuc. Acids Res. 10, 141-156 and Staden (1990) Meth. Enzymol. 183, 163-180).

Recombinant plasmid DNA is purified from Clone 2, linearised by cleavage of the unique EcoRI site and 5' capped RNA is transcribed using SP6 RNA polymerase and mCAP mRNA capping kit (Stratagene Inc.). In vitro translation reactions (25 µl) using 0.1µg RNA transcript, 20µCi [<sup>35</sup>S] methionine and a rabbit reticulocyte lysate system are carried out according to manufacturer's instructions (Promega Inc.). Confirmation that the SP6 RNA polymerase and the in vitro translation system are functional is obtained by running a positive control alongside. The control plasmid is a linearised SV64-carboxypeptidase E construct with an upstream SP6 promoter region (Fricke et al, (1989) Mol. Endocrinol. 3, 666-673).

A fragment of bacteriophage E DNA of 1892 bp containing the complete Clone 1 insert is excised from Clone 1 using EcoRI and XbaI. This is directionally cloned into the vector pGEM-IIz (Promega Inc.) cut with the same restriction enzymes thus placing a SacI site 3' of the Clone 1 insert. A 707bp SacI/AvrII fragment is excised from this new construct. This 707bp fragment encodes the predicted C-terminal 114 amino acids of the endosialidase and the 3' untranslated region of KIE DNA. It is ligated into the 3253 bp product of a SacI/AvrII digest of Clone 2. The resulting plasmid (Clone 3) contains only the extreme 5' and 3' regions of the originally cloned KIE DNA in a Bluescript SK+ vector effectively lacking the central 2975bp of the DNA sequence which includes the sequence encoding the predicted endosialidase open reading frame. A 2975bp fragment derived from an AvrII digest of total KIE DNA is ligated into Clone 3 digested with AvrII. The resulting construct in Bluescript SK+ (Clone 4) contains the full length endosialidase gene previously encoded in Clones 1 and 2, and the gene is sequenced using the Sequenase 2.0 sequencing kit (USB Corp). It has the sequence shown in SEQ, ID. No: 1.

Example 2:      Preparation of Recombinant Plasmid for Expression of  
Bacteriophage E Endosialidase

Clone 4, the DNA construct containing the complete endosialidase open reading frame prepared as described in Example 1, is subjected to PCR using primers 5'-CCGGGGATCCATGATTCAAAGACTAGGTTCTTCATTA-3' and 3'-CGTTAGACGACGTGCGGTCTTGTATCTTAAGACAC-5' to facilitate amplification of the endosialidase open reading frame with incorporation of a BamHI restriction site and an EcoRI restriction site at the 5' and 3' termini of the open reading frame respectively.

The 2483 bp PCR product is cleaned by extraction first with a mixture of equal volumes of phenol (equilibrated to pH 8.0 with 2-amino-2-hydroxymethylpropane-1,3-diol) and a 24:1 mixture of chloroform and isoamyl alcohol, then with the chloroform: isoamyl alcohol mixture alone, followed by precipitation in ethanol and resuspension in TE buffer (10 mM 2-amino-2-hydroxymethylpropane-1,3-diol hydrochloride, 1 mM EDTA, pH8.0). The cleaned PCR product and pGEX-2T expression vector (Pharmacia Biotech) are digested simultaneously with BamHI and EcoRI and purified by agarose gel electrophoresis and Qiaex extraction (Quiagen Corp). The cut PCR product and expression vector are ligated using T4 DNA ligase (New England Biolabs) according to the manufacturer's instructions, to form a recombinant vector, which is sequenced, using USB Sequenase 2.0, across the two cloning sites to verify that the correct reading frame has been maintained.

Example 3:      Transformation and Expression

The ligation product of Example 2 is used to transform electrocompetent *E. coli* MC 1061 cells using a Bio Rad electroporation apparatus and the transformed cells are selected by restriction analysis of a miniprep DNA sample obtained from the cells.

The transformed cells are cultured to express a fusion protein by the addition of IPTG to a final concentration of 0.5mM, when the OD<sub>600</sub> has reached 0.3, and are then allowed to grow for a further 4 hours at 37°C with shaking at 250 rpm. The expressed fusion protein is purified from the culture medium by binding to glutathione beads as described by Smith and Johnson (1988) Gene 67, 31-40.

Samples of the bacterial culture and of purified fusion protein fractions are subjected to SDS-PAGE, electrophoretically transferred to a nitrocellulose membrane, washed and hybridised to antiendosialidase polyclonal antiserum by the method of Sambrook et al, (1989), op. cit. Immunoreactive bands are detected by binding a second antibody conjugated to alkaline phosphatase and reaction with (a) a 50mg/ml solution of nitroblue tetrazolium chloride in a 70:30 mixture of dimethylformamide and water and (b) a 50mg/ml solution of 5-bromo-4-chloro-3-indolyl phosphate disodium salt in water.

The release of N-acetylneuraminic acid (NANA) from polysialic acid by purified fractions of the fusion protein is measured using the TBA assay of Horgan (1981) Clin. Chim. Acta 116, 409-415. The measurements show that rate of release of NANA is directly proportional to fusion protein concentration. No release of NANA is observed when the fusion protein is replaced by glutathione S-transferase protein alone.

- 13 -

SEQUENCE LISTING

(I) INFORMATION FOR SEQ. ID. NO: 1

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2436 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ORIGINAL SOURCE:

(A) ORGANISM: Bacteriophage E

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1 .. 2436

(D) OTHER INFORMATION: / product = "coding region for bacteriophage E endosialidase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1

-14-

ATGATCAGAACAGATA GCTTCTCATTTATT ATTACGAGCACTAA ATACAGCTCCTTC TGCGTACTCTGGT GACAGCTCACCG CTTGATCCGTTAA GATTGGGCC  
 G C D G K R T H D P D A V N A M A S C K R I D G A G A T T Y K V \$ S L P D H E R F  
 GCTGATGTCAGACA AACGACAGAGATCA CTAACTGCAGAGTC GATTCAGCAGGAG GCTTCAGGAGCA ATTGGCGCTGT GETTCAGACAGA TCATCTTACCTCT ATGGAGCGATT  
 N T R F V H E R L A G O P L Y Y U S K C G E L D K L T E V V S L N D F G A  
 MACACCGCTCTGA TGGGACGTTTACA GCTAACCTCTTC TATGGCTGAGCT TTATCAGTGGA CTCATTAATACG GATTCGACCTAC ANTCTGGCT  
 D K A F V I E H V I Y A P Y M G S D A H G V S R I H V S M V K S G D D G O T W  
 GACAGAGCTGTA TATGGAGCTTA TATGGCTCTAGT GTTGGCCACCTCT GCTTGTACCTCTG CTGATCTGGT AACGCTGCTAC GCTGACAGT  
 T P F M I T D H II P D Y P T V N Y H C M S M G V C R N R I F A M I E T R A T L A  
 ACTGGAGCTTA ACTGGAGCTTA GCTAACCTCTTC AACATCTCTG AGCTGGCTGAGT AGCTGGCTGAGT CCACACCTCTGT GCTATGAGACAA CCTACTTGGCC  
 N E L T M C A L M D H M S L H T G C G I T K A A N O R Y A T I H V P D H  
 AACGAGCTCCAT TETGGATCTGGAT CCCTCTATCTCT AGTGCCTACTAGT GCTGAGCTGAGA TATGAGCTACCT GCTCTGATAC  
 L P V G D F V N F S N S A V T G V S G D M M K V A T V I H V L T P N O  
 CTCCTCTGGAT TITGTTACTCTCT AACCTGGGAGAA CCTGAGCTGAGT AACGAGCTGAGCA GNTAGAGTAAAG AACCTGGAGCT  
 T S D L N N A C K N W H M G T S F II K S P M R K T D L G L I P A V T E V H S F  
 ACTTCAGTCTGAT AACGGTGERAGAAAT TGGCATGGCTACT TCTTCCATGACTCT CGCTGGCTAGACA GATCTGCTCCTAC CTCGGCTCAG  
 T D I N N G F V H G D V A P A E V G L F Y P D A F N S P S N Y V R R L  
 TATGATGAGAT GCTTGTATGGC TATCATGGCTGT GNGCTCAGAGAA GTGGCTTTCAC TCTCTGAGCTTC ATAGCCCTCTAT TATCTCTGG  
 I P S E Y E P D A A E P C J K Y D G V L Y L I T R A G D R L G S S L H R  
 ATACATGGAT GACCAAGCCGA GACCATCAGCAG TACGGAGCTGTA TATGGCTACT GCTGACTCTGC GACGAGCTGGAC TCTCCTGGT  
 R D I G D T M E S L R F P H N V H H T T L P F A N V C D D L I M F G S E R A A E  
 AGCTGATAGCTG ACTGGAGCTCA AGATTCACATCT CTGGATCATACT TACCTGGCTCTAG CTGAGCTACCT ATTATCTGCTCA GACCTGGAG  
 E M A P D D R A K A S Y P R T F Y A R L N V N N N N A D D I E M V N I L T  
 GATGGGAGAGT GCACTGAGCTG TACGGAGCTCTT GCTGACCTCTT GCACTGAGCTG AACATGGAGCTCA AACGAGCTGAGCA  
 0 1 Y 0 G D I V H S S V G V C S V W K D S F I Y I F G G E N H F N P M T  
 CAACTCTGGT GACATGCTGACTCT AGTGGAGCTGACTCT TCTGTTAGTTAA GACGAGCTCTTC TATCTCTGGT GATGGAGCTCTT ATTATCTGCTCA GACCTGGAG  
 D N K D K D P F R K G H G H P T D I Y C Y K M O I A N N R V A P L F T Y G A  
 GACACAGACGAA GACCATTTAGGT CATGAGACCTCTT GATAGAGCTCTT AACATGGAGCTCA AACGAGCTGAGCA AACGAGCTGAGCT  
 G O A I P T F H G T D G I R N I P A P L Y F S D N I V T E D I K V G H N L P  
 GCTGAGCTGACTCT ACTCTGGCTGACTCT GATGGAGCTGACTCT AACCTGGCTCTAGT GATGGAGCTCTT GATGGAGCTCTT ATTATCTGCTCA  
 S T S A N I R S E M O M E G E Y G F I G K S V P K D K P T G O R L I I C G E  
 AACGAGCTGCTT ATTACGAGCTG CAGTGGAGCTG TATGGTTATGGC AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT  
 T S S S S G A Q O I T L H G S N S S N H T Y N G N E H L P Q G A P I M P A  
 ATCTGAGCTCTT GCTGAGCTGACTCT TCTGAGCTCTAGT GCTATCTGAGCTG GATGGAGCTGACTCT AACGAGCTGAGCT AACGAGCTGAGCT  
 D N O P A A G G P S N R F T T I Y I G S D P V T T S D A D H I S I S N T  
 GATGGAGCTGACTCT GCTGAGCTGACTCT AACGAGCTGACTCT AACCTGGCTCTAGT GACCTGGCTCT AACGAGCTGAGCT AACGAGCTGAGCT  
 V L R K A M S R V G F K Q Y G L N S E A E R N L D S I H F G V L A O D I V A H F  
 GCTGAGCTGACTCT AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT  
 A E G L A I K Y G I V S F E E G R Y G V R Y S E V L I L E A A Y T R H A L D  
 GCTGAGCTGACTCT GCTGAGCTGACTCT AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT AACGAGCTGAGCT  
 L E E M Y A T N K I S

Claims

1. A recombinant protein having bacteriophage endosialidase enzymatic activity obtainable by expression from a recombinant vector comprising a DNA sequence encoding a bacteriophage endosialidase linked to a DNA sequence of an expression vector which expresses a polypeptide which adds to the N- terminus of the endosialidase, or an analogue of said protein which is a mutant, functional fragment or derivative of said protein having endosialidase enzymatic activity.
2. A protein or analogue according to claim 1, in which the DNA sequence encoding the endosialidase is derived from a DNA construct encoding amino acid residues encoded by nucleotides 172 to 1744 of SEQ. ID. No. 1.
3. A protein or analogue according to claim 1, in which the DNA sequence encoding the endosialidase is derived form a DNA construct encoding amino acids encoded by nucleotides 1 to 2436 of SEQ. ID. No. 1.
4. A protein according to claim 1, 2 or 3 which is a fusion protein comprising the endosialidase linked directly or through a spacer to a polypeptide derived from the expression vector, or an analogue of said protein which is a mutant, functional fragment or derivative of said protein.
5. A protein or analogue according to claim 4 in which the polypeptide is glutathione S-transferase.
6. A protein or analogue according to any of claims 1-5, in which the DNA sequence encoding the endosialidase is derived from a DNA construct comprising a DNA fragment encoding the endosialidase cloned into a bacterial cloning vector.
7. A protein or analogue according to claim 6, in which the DNA fragment is derived from genomic bacteriophage E DNA or a synthetic DNA which is substantially homologous thereto.

8. A protein or analogue according to claim 6 or 7, in which the DNA construct is amplified by a polymerase chain reaction using primers incorporating restriction sites.
9. A recombinant vector comprising a DNA sequence encoding a bacteriophage endosialidase linked to a DNA sequence of an expression vector which expresses a polypeptide which adds to the N- terminus of the endosialidase, said recombinant vector being capable of directing expression of said protein in a compatible host cell.
10. A recombinant vector according to claim 9 in which said encoding DNA sequence is a DNA construct according to any of claims 6 to 8.
11. A recombinant vector according to claim 9 or 10, in which the expression vector is a prokaryotic expression vector.
12. A recombinant vector according to any of claims 9 to 11, in which the expression vector is a pGEX vector.
13. A recombinant vector according to any of claims 9 to 12, in which the expression vector is pGEX - 2T.
14. A process for the production of a protein according to any of claims 1 to 8 which comprises culturing a host cell transformed with a recombinant vector according to any of claims 9 to 13 under conditions allowing expression of said protein, and isolating the protein thereby produced.
15. A process according to claim 14, in which the host cell is a microbial cell.
16. A process according to claim 14, in which the host cell is an E. coli cell.
17. A host cell transformed with a recombinant vector according to any of claims 9 to 13.
18. A host cell according to claim 17, which is a transformed microbial cell.

-17-

19. A host cell according to claim 17 which is a transformed *E. coli* cell.
20. A pharmaceutical composition comprising a protein or analogue according to any of claims 1 to 8, or a pharmaceutically acceptable salt thereof, optionally together with a physiologically acceptable carrier.
21. A protein or analogue according to any of claims 1 to 8 for use in the diagnosis or treatment of a medical condition.
22. Use of a protein or analogue according to any of claims 1 to 8 in the preparation of a medicament for the diagnosis or treatment of meningitis or other condition caused by *E. coli* K1 or by other bacteria expressing polysialic acid on the cell surface thereof, of cancer characterised by expression of polysialic acid on the surface of a tumour cell or of tumour metastasis.
23. Use of a protein or analogue according to any of claims 1 to 8 in the analysis of a glycoprotein.
24. A method of treating a condition caused by a bacterium expressing polysialic acid on a cell surface thereof, cancer characterised by expression of polysialic acid on a tumour cell surface, or tumour metastasis, which comprises administering a protein or analogue according to claim 1 to a warm-blooded mammal in need of such treatment.

**INTERNATIONAL SEARCH REPORT**

International Application No  
PL 1/GB 96/01577

A. CLASSIFICATION OF SUBJECT MATTER					
IPC 6	C12N15/70	C12N15/62	C12N9/24	C12N1/21	A61K38/47

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
**IPC 6 C12N**

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MOL. MICROBIOL., vol. 16, no. 3, May 1995, pages 441-450, XP000604798 GERARDY-SCHAHN ET AL.: "Molecular cloning and functional expression of bacteriophage PK1E-encoded endoneuraminidase Endo NE" see page 443 - page 445 see page 449	1-4, 6-11, 14-19,21
Y	---	5,12,13
A	---	20,22-24
Y	EP,A,0 587 541 (TECNOGEN S.C.P.A.) 16 March 1994 see page 4; examples 1,2 ---	1,4,5, 8-21
	-/-	

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- 'A' document defining the general state of the art which is not considered to be of particular relevance
- 'E' earlier document but published on or after the international filing date
- 'L' document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- 'O' document referring to an oral disclosure, use, exhibition or other means
- 'P' document published prior to the international filing date but later than the priority date claimed

- 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- '&' document member of the same patent family

Date of the actual completion of the international search  14 October 1996	Date of mailing of the international search report  28.10.96
Name and mailing address of the ISA  European Patent Office, P.B. 5818 Patenttaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Authorized officer  Gac, G

## INTERNATIONAL SEARCH REPORT

Int'l Application No  
PCT/GB 96/01577

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,93 23544 (THE UNIVERSITY COURT OF THE UNIVERSITY OF GLASGOW) 25 November 1993 see abstract see page 5 ---	1,4,5, 8-21
Y	EP,A,0 293 249 (AMRAD CORPORATION LIMITED) 30 November 1988 see whole document, especially Figure 7 and claims 2, 3 ---	1,4,5, 9-21
Y	EP,A,0 646 646 (BRISTOL-MYERS SQUIBB COMPANY) 5 April 1995 see the whole document -----	1,4,5, 9-21

1

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB96/01577

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: 24 because they relate to subject matter not required to be searched by this Authority, namely.  
Remark: Although claim 24 is directed to a method of treatment of (diagnostic method practised on) the human/animal body((Art.17.2.a.i) and (Rule 39.2.iv)PCT)), the search has been carried out and based on the alleged effects of the compound/composition.
2.  Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

#### Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International Application No	
PCT/GB 96/01577	

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-587541	16-03-94	IT-B-	1262981	23-07-96
W0-A-9323544	25-11-93	AU-B- AU-A- CA-A- EP-A- JP-T- US-A-	666013 4080493 2113437 0607368 6509237 5482708	25-01-96 13-12-93 25-11-93 27-07-94 20-10-94 09-01-96
EP-A-293249	30-11-88	AU-B- AU-A- WO-A- DE-A- DK-B- ES-T- JP-B- JP-T- NO-B-	607511 1793288 8809372 3873989 169975 2045115 6081596 1503441 178894	07-03-91 21-12-88 01-12-88 01-10-92 18-04-95 16-01-94 19-10-94 22-11-89 18-03-96
EP-A-646646	05-04-95	CA-A- JP-A-	2132937 7184663	31-03-95 25-07-95